Draft genome sequence of a novel Bacillus glycinifermentans strain having antifungal and antibacterial properties

A. Karim, Olivier Poirot, A. Khatoon, M. Aurongzeb

Draft genome sequence of a novel *Bacillus glycinifermentans* strain having antifungal and antibacterial properties

A. Karim[a,*], O. Poirot[b], A. Khatoon[a], M. Aurongzeb[a]

[a] *Jamil-Ur-Rahman Center for Genome Research, Dr Panjwani Center for Molecular Medicine and Drug Research (PCMD), International Center for Chemical and Biological Sciences (ICCBS), University of Karachi, Karachi 75270, Pakistan*

[b] *Aix-Marseille Université, CNRS, Structural and Genomic Information Laboratory, UMR 7256 (IMM FR 3479), 163 Avenue de Luminy, Case 934, 13288, Marseille Cedex 9, France*

Objectives: Bacillus spp. have been used as biocontrol agents against soilborne pathogens because they produce secondary metabolites that exhibit a wide range of antibacterial or antifungal properties. In this study, a novel strain of *Bacillus glycinifermentans* sp. (JRCGR-1) was identified and its genome was sequenced and annotated. The genome was explored for putative genes involved in antimicrobial activity.
Methods: Whole-genome sequencing was performed on an Illumina NextSeq 500 platform. Read quality was checked by FastQC, paired-end reads were trimmed using Sickle, and de novo assembly was performed using SPAdes v.3.11.11. QUAST 5.02 was used to assess the quality of contigs and scaffolds. Finally, the assembled scaffolds were annotated by Prokka v.1.13. Genes involved in antimicrobial metabolite biosynthesis were predicted using antiSMASH. Virulence and antimicrobial resistance genes were predicted using BacWGSTdb and the Comprehensive Antibiotic Resistance Database (CARD), respectively.
Results: The genome of *B. glycinifermentans* JRCGR-1 was 4 700 692 bp in size with a G + C content of 45.52%. Final assembly of the genome resulted into 84 contigs and 83 scaffolds (>500 bp length). Overall, the genome comprises 5174 genes, 32 tRNAs, 4 rRNAs, 1 tmRNA and 92 misc_RNAs. Eleven putative gene clusters responsible for antimicrobial metabolite biosynthesis were identified, including genes for biosynthesis of non-ribosomal lipopeptides and polyketides. Virulence and antimicrobial resistance genes were also identified in the genome.
Conclusion: The presence of antimicrobial resistance genes in the genome of *B. glycinifermentans* JRCGR-1 makes it a potential biocontrol agent against soilborne pathogens.

Soilborne pathogens are emerging as a global threat to agricultural products, and each year worldwide they cause huge economic losses. Several chemical-based fertilisers and pesticides are used to control soilborne pathogens. However, with long-term use these chemicals are hazardous to human health and cause environment pollution. An alternative to overcome this problem is the use of biocontrol agents [1]. Specifically, the genus *Bacillus* is a potent candidate for biocontrol agents against soilborne pathogens as members of this genus produce a wide range of secondary metabolites, hydrolytic enzymes and volatile organic compounds with antibacterial or antifungal properties.

Much work has been done to explore the biochemistry and physiology of *Bacillus* spp. Advancements in the field of whole-genome sequencing have given a new dimension to microbial-based products and processes.

In 2015, two bacterial isolates from cheonggukjang, a Korean fermented soybean paste food product, were found to comprise a novel *Bacillus* sp. named *Bacillus glycinifermentans* sp. nov. [2]. To date, only five *B. glycinifermentans* genome sequences are available in public databases, including three complete genomes (GenBank Assembly accession nos. **GCA_900093775.1**, **GCA_002443095.1** and **GCA_004103615.1**) and two draft genomes (GenBank Assembly accession nos. **LECV00000000.1** and **LECW00000000.2**). To the best of our knowledge, this is the first draft genome sequence of *B. glycinifermentans* reported from Pakistan.

To study the genome of *B. glycinifermentans* strain JRCGR-1, the strain was first cultured on a nutrient agar plate. Colonies were then inoculated into 10 mL of Luria broth and were incubated at 37 °C for 24 h. The optical density was adjusted to a McFarland

standard of 3–4nm. Following incubation, bacterial cells were collected by centrifugation (3000 × *g* for 15 min). The supernatant was removed without disturbing the sediment and the DNA was extracted using a commercial DNA extraction kit (QIAamp® DNA Mini Kit; QIAGEN, Hilden, Germany). A Qubit™ dsDNA BR Assay kit (Invitrogen; Thermo Fisher Scientific, Eugen, OR, USA) was used to calculate the amount of DNA using a Qubit® 2.0 fluorometer (Invitrogen) according to the manufacturer's instructions. A Nextera XT DNA Library Kit (Illumina Inc., San Diego, CA, USA) was used to construct a paired-end library, and sequencing was performed on a NextSeq 500 platform (Illumina Inc.) in paired-end read mode. The quality of the reads was checked using FastQC software. Paired-end reads were then trimmed using Sickle and de novo assembly was performed using SPAdes v.3.11.11 into contigs and scaffolds (base

Table 1
Secondary metabolite profile of six *Bacillus glycinifermentans* strains.

Strain/metabolite type	Most similar known cluster	Similarity (%)
B. glycinifermentans GO-13		
Lanthipeptide	Geobacillin II	50
NRPS	Lichenysin	100
T3PKS	–	–
NRPS	Fengycin	26
Siderophore	–	–
β-Lactone	Fengycin	53
Lasso peptide	–	–
NRPS	Bacillibactin	53
NRPs, T1PKS, terpene	Paenibacterin	60
B. glycinifermentans KJ-17		
Thiopeptide, bacteriocin	Butirosin	7
NRPS, terpene, T1PKS	Paenibacterin	60
Lasso peptide	–	–
T3PKS	–	–
Lanthipeptide	Geobacillin	50
Siderophore	–	–
NRPS	Fengycin	33
NRPS	Bacillibactin	46
NRPS	Lichenysin	100
β-Lactone	Fengycin	53
B. glycinifermentans BGLY		
Sactipeptide	Sporulation killing factor SkfA	85
NRPS	Lichenysin	100
Thiopeptide, bacteriocin	Butirosin	7
Siderophore	–	–
β-Lactone	Fengycin	53
Terpene	–	–
T3PKS	–	–
NRPS	Bacitracin	88
NRPS	Bacillibactin	53
B. glycinifermentans KBNO06PO3352		
NRPS	Fengycin	26
β-Lactone	Fengycin	53
NRPS, T1PKS, terpene	Paenibacterin	60
T3PKS	–	–
NRPS	–	–
Lasso peptide	–	–
Lanthipeptide	Geobacillin	50
NRPS	Lichenysin	100
Thiopeptide, bacteriocin	Butirosin	7
Siderophore		
B. glycinifermentans SRCM103574		
Sactipeptide (head-to-tail cyclised peptide)	Sporulation killing factor SkfA	85
NRPS	Lichenysin	100
Thiopeptide, bacteriocin	Butirosin	7
Siderophore	–	–
β-Lactone	Fengycin	53
Terpene	–	–
T3PKS	–	–
NRPS	Bacitracin	88
NRPS	Bacillibactin	53
B. glycinifermentans JRCGR-1		
T3PKS	–	–
Terpene	–	–
β-Lactone	Fengycin	53
NRPS	Lichenysin	100
Head-to-tail	Sporulation killing factor SkfA	71
Siderophore	–	–
NRPS	Bacillibactin	53
NRPS	Bacitracin	66
Bacteriocin	–	–
NRPS	–	–
NRPS	Bacitracin	33

NRPS, non-ribosomal peptide synthetase.

quality scores > Q20, $k = 39$). The quality of contigs and scaffolds was evaluated by QUAST 5.02. The assembled contigs were annotated using Prokka v.1.13 [>500 bp; e-value cut-off default (10_6)] for rapid annotation of prokaryotic genomes. tRNAscan-SE was used for the prediction of tRNA genes. The plasmid was assembled using plasmidSPAdes (base quality scores > Q20, $k = 55$), and gene annotation was done using Prokka v.1.13 [>500 bp; e-value cut-off default (10_6)]. RNAmmer and Barrnap were used to identify RNA genes. Putative genes involved in antimicrobial metabolite biosynthesis were predicted using antiSMASH v.5.0.0rc1. Virulence genes were predicted using the BacWGSTdb service (http://bacdb.org/BacWGSTdb/index.php). Antimicrobial resistance genes were predicted using the Comprehensive Antibiotic Resistance Database (CARD) (https://card.mcmaster.ca/home).

A total of 4 851 845 paired-end reads (2×76-bp) were generated with $32\times$ coverage for strain JRCGR-1. The genome assembly contained 4 700 692 bp and with an average G + C content of 45.52%. The final assembly contains 84 contigs and 83 scaffolds (>500 bp length, N_{50} of 135 232 bp). The maximum contig size was 384 553 bp. Overall, the genome comprises 5174 genes, 32 tRNAs, 4 rRNAs, 1 tmRNA and 92 misc_RNAs. The plasmid was assembled into 37 scaffolds using plasmidSPAdes. Other features of the plasmid include a size of 1 113 267 bp, 27 tRNAs, 4 rRNA, 1366 genes, 21 misc_RNAs and 1314 CDS (coding sequences).

Genes involved in antimicrobial metabolite biosynthesis were predicted using antiSMASH. Draft genome analysis of strain JRCGR-1 revealed 11 putative gene clusters responsible for antimicrobial metabolite biosynthesis, among which five encode non-ribosomal peptide synthetases (NRPS) (for biosynthesis of one fengycin, one lichenysin, one bacillibactin and two bacitracins) and one related to terpene and β-lactone biosynthesis. Siderophores (bacillibactin) produced by bacteria are involved in inhibition of phytopathogen growth by depriving them of essential iron [3].

Supplementary Figs. S1 and S2 show two putative gene clusters of strain JRCGR-1 (label as query sequence) involved in biosynthesis of secondary metabolites. These genes have similarities with non-ribosomal lipopeptides (lichenysin, surfactin, fengycin, mycosubtilin and plipastatin) and polyketides (basiliskamides). Supplementary Fig. S1(A) shows the first putative gene cluster of strain JRCGR-1 involved in secondary metabolite biosynthesis and its similarity to lichenysin (100%), surfactin (47%) and basiliskamides (9%). Comparison of this gene against other *Bacillus* spp. shows 94% similarity with *B. glycinifermentans* strain GO-13, *B. glycinifermentans* strain KBNO06PO3352 and *B. glycinifermentans* strain BGLY [Supplementary Fig. S1(B)]. Similarly, the second putative gene cluster presented similarities with fengycin (53% similarity), mycosubtilin (40% similarity) and plipastatin (30% similarity) [Supplementary Fig. S2(A)]. Comparison of this predicted gene cluster of strain JRCGR-1 with other *Bacillus* spp. showed similarity between 45–50% for seven *Bacillus paralicheniformis* sp. and 46% gene similarity for *Bacillus licheniformis* strain B4123 [Fig. S2(B)]. Secondary metabolites for five other strains of *B. glycinifermentans* were also predicted using antiSMASH. The finding reveals that fengycin and lichenysin were present as core genes, i.e. they were present in all six strains. A putative gene for bacillibactin was absent in *B. glycinifermentans* strain KBNO06PO3352. Sporulation killing factor SkfA was present in *B. glycinifermentans* strain JRCGR-1 and *B. glycinifermentans* strain SRCM103574. Details of the secondary metabolite profile of six strains of *B. glycinifermentans* sp. are given in Table 1. The presence of these NRPS genes can be linked with the biocontrol potential of bacteria [4]. Several gene cluster were also identified that can suppress the growth of Gram-positive bacteria by producing synthetases (PKS), e.g. oleandomycin, β-lactamase, tetracenomycin and tetracycline. Putative *bsIA* (*yuaB*) and *tasA* genes associated with the production of biofilm matrix and fungal cell-wall-degrading enzymes, respectively, were also identified. In the genomes of six strains of *B. glycinifermentans* sp., two antimicrobial resistance genes (*bcrA* and *Escherichia coli ampC1* β-lactamase) were present as accessory genes. These genes were present in three strains of *B. licheniformis*, including SRCM103574, BGLY and JRCGR-1. The *bcrA* gene is an ATP-binding cassette (ABC) transporter found mostly in *B. licheniformis* sp. that confers bacitracin resistance [5]. The *E. coli* AmpC1 β-lactamase resistance mechanism is based on antibiotic inactivation, whereas for BcrA the resistance mechanism involves antibiotic efflux. In the genomes of strains KBNO06PO3352 and GO-13, only *E. coli ampC1* β-lactamase was predicted. On the other hand, strain KJ-17 lacks all of these genes. We also identified virulence genes in the strain JRCGR-1 genome, including *capA*, *capB*, *capC* and *hlyIII*, with similarities of 83.25%, 89.30%, 88% and 82.26%, respectively.

The aforementioned findings reveal that *B. glycinifermentans* JRCGR-1 might be a potential candidate as a biological control agent against plant soilborne diseases.

GenBank accession no

The draft genome sequence of *B. glycinifermentans* JRCGR-1 was submitted to the NCBI database with accession no. **VHPY0000000**.

Funding

This study was supported by funds from the International Center for Chemical and Biological Sciences (ICCBS), University of Karachi (Karachi, Pakistan), Campus France and Embassy of France in Pakistan.

Competing interests

None declared.

Ethical approval

Not required.

Acknowledgments

The authors would like to thank Matthieu Legendre, Chantal Abergel, Sébastien Santini and Sébastien Nin [Aix-Marseille Université, CNRS, Structural and Genomic Information Laboratory, UMR 7256 (IMM FR 3479)] and Georges Massiot [University of Reims, Reims, France] for their valuable support.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jgar.2019.10.011.

References

[1] Cook RJ. Making greater use of introduced microorganisms for biological control of plant pathogens. Annu Rev Phytopathol 1993;31:53–80.

[2] Kim S-J, Dunlap CA, Kwon S-W, Rooney AP. *Bacillus glycinifermentans* sp. nov., isolated from fermented soybean paste. Int J Syst Evol Microbiol 2015;65:3586–90.

[3] Chen XH, Koumoutsi A, Scholz R, Eisenreich A, Schneider K, Heinemeyer I, et al. Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. Nat Biotechnol 2007;25:1007–14.

[4] Zhang S, Jiang W, Li J, Meng L, Cao X, Hu J, et al. Whole genome shotgun sequence of *Bacillus amyloliquefaciens* TF28, a biocontrol entophytic bacterium. Stand Genomic Sci 2016;11:73.

[5] Podlesek Z, Comino A, Herzog-Velikonja B, Žgur-Bertok D, Komel R, Grabnar M. *Bacillus licheniformis* bacitracin-resistance ABC transporter: relationship to mammalian multidrug resistance. Mol Biol 1995;16:969–76.